Estimating long term behavior of flows without trajectory integration: the infinitesimal generator approach

Gary Froyland*   Oliver Junge†   Péter Koltai‡

January 24, 2011

Abstract

The long-term distributions of trajectories of a flow are described by invariant densities, i.e. fixed points of an associated transfer operator. In addition, global slowly mixing structures, such as almost-invariant sets, which partition phase space into regions that are almost dynamically disconnected, can also be identified by certain eigenfunctions of this operator. Indeed, these structures are often hard to obtain by brute-force trajectory-based analyses. In a wide variety of applications, transfer operators have proven to be very efficient tools for an analysis of the global behavior of a dynamical system.

The computationally most expensive step in the construction of an approximate transfer operator is the numerical integration of many short term trajectories. In this paper, we propose to directly work with the *infinitesimal generator* instead of the operator, completely avoiding trajectory integration. We propose two different discretization schemes; a cell based discretization and a spectral collocation approach. Convergence can be shown in certain circumstances. We demonstrate numerically that our approach is much more efficient than the operator approach, sometimes by several orders of magnitude.

1 Introduction

Analysis of the long-term behavior of flows can be broadly classified into geometric methods and statistical methods. Geometrical methods include the determination of fixed points, periodic orbits, and invariant manifolds. Invariant manifolds of fixed points or periodic orbits act as barriers to transport as trajectories may not cross the manifolds transversally. Statistical methods include determining the distribution of points in very long trajectories of a very large set of initial points (i.e. a physical invariant measure [29], often possessing an invariant density) and the identification of meta-stable or almost-invariant sets [8, 15, 14]. Almost-invariant sets partition the phase space into almost dynamically disconnected regions and are important for revealing global dynamical structures that are often invisible to an analysis of trajectories. These metastable dynamics also go under the names of persistent patterns, or strange eigenmodes, both of which are realisable as eigenfunctions of a transfer operator. Frequently, the boundaries of maximally almost-invariant sets (those sets which are locally closest to invariant sets) coincide with certain invariant manifolds [16].

Our focus in the present paper is on statistical methods, although we demonstrate via case studies the relationships to geometrical methods. The most commonly used tool for statistical methods is the transfer operator (or Perron-Frobenius operator). Fixed points of the transfer operator correspond to invariant densities, while eigenfunctions corresponding to real positive eigenvalues strictly less than one provide information on almost-invariant sets. In practice, one typically constructs a finite-rank numerical approximation of a transfer operator and computes

---

*School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.

†Faculty of Mathematics, Technische Universität München, 85748 Garching, Germany.

‡Faculty of Mathematics, Technische Universität München, 85748 Garching, Germany. Péter Koltai was partly supported by the TopMath PhD program within the Elite Network of Bavaria and the TUM Graduate School.

large spectral values and eigenfunctions for this finite-rank operator. The construction of the finite-rank approximation requires the integration of many relatively short trajectories with initial points sampled over the domain of the flow. It is this use of *short* trajectories that gives the transfer operator approach additional stability and accuracy when compared with computations based upon very long trajectories. Long trajectories continually accumulate small errors from imperfect numerical integration and finite computer representation of numbers; these small errors quickly grow in chaotic flows. While the transfer operator approach is very stable, it still requires the computation of many small trajectories which can be very time consuming in some systems. The approach we describe in the present work obviates the need for any trajectory integration at all and works directly with the vector field.

Our approach exploits the fact that the evolution of probability densities $u = u(t)$ can be described by generalized solutions of the abstract Cauchy problem $\partial_t u = \mathcal{A}u$. The Perron–Frobenius operator is the evolution operator of this equation, and has the same eigenfunctions as the operator $\mathcal{A}$. The operator $\mathcal{A}$ is an unbounded hyperbolic (if the underlying dynamics is deterministic) or elliptic (if the deterministic dynamics is perturbed by white noise) partial differential operator. Standard techniques allow us to approximate the eigenmodes we are interested in: finite difference, finite volume, finite element and spectral methods yield such discretizations; see [25, 19, 21, 4, 5] and the references therein.

An outline of the paper is as follows. In Section 2 we provide background on the infinitesimal operator arising from smooth vector fields and describe conditions under which the operator generates a semigroup of transfer operators. In order to obtain formal results, we will require the addition of a small amount of diffusion to the deterministic flow. In the small diffusion setting we discuss existence of invariant densities and spectral results for the associated infinitesimal operator. Section 3 describes a spectral Galerkin method for the approximation of the infinitesimal operator. We apply results from the numerical analysis of advection-diffusion PDEs to show that our Galerkin method approximates the true eigenfunctions of the infinitesimal operator as our Galerkin basis becomes increasingly refined. We can also show that the convergence rate is spectral; that is, faster than any polynomial. Section 4 describes an Ulam-based Galerkin method for approximating the infinitesimal operator. This Ulam-based approach is new and shares some similarities with finite-difference schemes. While we cannot show convergence of this approximation scheme, the numerical results obtained are extremely fast and accurate. Sections 5–8 detail the practical application of our two approximation methods to flows in one-, two-, and three-dimensional domains. We demonstrate the spectral accuracy of our spectral Galerkin method and compare with the accuracy of the Ulam-based Galerkin method. We also compare the accuracy vs. computational effort of our two new approaches with standard transfer operator approaches. The natural relationships between the outputs of our statistical methods and geometric objects such as the vector field and invariant manifolds are also elucidated in each case study. We conclude in Section 9.

2 Dynamics, densities and semigroups

Let the domain $M \subset \mathbb{R}^d$ of our flow be a smooth compact manifold and m the (normalized) Lebesgue measure on M. Denote by $F : M \to \mathbb{R}^d$ the vector field generating the flow and by $\Phi^t : M \to M$, $t \in \mathbb{R}$, the flow, i.e. $\Phi^t(x)$ represents the location of a trajectory beginning at $x \in M$ after flowing for $t \in \mathbb{R}$ time units. One has that $d\Phi^t(x)/dt = F(x)$. Note that – provided that the components $F_i, i = 1, \ldots, d$ have continuous derivatives – the function $\Phi^t : M \to M$ is a diffeomorphism for every $t \in \mathbb{R}$.

Invariant sets are structures of dynamical interest. A set $A \subset M$ is called *invariant*, iff $\Phi^{-t} A = A$ for all t. Also, one asks how the flow changes probability measures. Sample x according to a probability measure μ; the distribution of $\Phi^t x$ is then given by $\mu \circ \Phi^{-t}$. Special attention is to be drawn to *invariant measures*, which do not change under the dynamics ($\mu = \mu \circ \Phi^{-t}$). Invariant measures μ are called *ergodic* if invariant sets have either zero or full measure, i.e. if $A \subset M$ satisfies $\Phi^{-t} A = A$ then $\mu(A) \in \{0, 1\}$. An even more restricted class of ergodic measures

are the *physically relevant* (or *natural*) ones, satisfying

$$\int \psi \, \mathrm{d}\mu = \lim_{T \to \infty} \frac{1}{T} \int_0^T \psi \left(\Phi^t x \right) \, \mathrm{d}t$$

for all $\psi : M \to \mathbb{R}$ continuous and $x \in U \subset M$ with $m(U) > 0$. One has that absolutely continuous[1] ergodic measures are natural invariant measures. The density of an invariant measure is called the *invariant density*.

2.1 Transfer operators

When looking for invariant densities, one can rephrase the action of the flow on measures as an action on densities. If f denotes the density of μ, then f is evolved by the flow as

$$\mathcal{P}^t f \, (x) = f \left(\Phi^{-t} x \right) \left| D\Phi^{-t} x \right|,$$

where $|B|$ denotes $|\det B|$ for a matrix B. The linear operator $\mathcal{P}^t : L^1(M) \circlearrowleft$ is known as a *transfer operator* or the *Perron-Frobenius operator* associated with the flow Φ. Note that invariant densities are fixed points of $\mathcal{P}^t$. For each $t \in \mathbb{R}$ we have

  (i) $\mathcal{P}^t$ is linear,

  (ii) $\mathcal{P}^t f \geq 0$ if $f \geq 0$ and,

  (iii) $\|\mathcal{P}^t f\| = \|f\|$ for all $f \geq 0$, where $\| \cdot \|$ is the L^1 norm.

i.e. $\mathcal{P}^t$ is a *Markov operator*. Moreover, these properties also imply $\|\mathcal{P}^t\| \leq 1$, hence $\mathcal{P}^t$ is a contraction and its eigenvalues lie inside the unit disc. In fact, since the flow Φ^t is a bijection, we have $\|\mathcal{P}^t f\| = \|f\|$ for all $f \in L^1$. To see this, note that

$$\int_A \mathcal{P}^t f \, dm = \int_{\Phi^{-t} A} f \, dm \tag{1}$$

for $t > 0$ and $A \subset M$ measurable. Now write $f = f^+ - f^-$, with $f^\pm(x) = \max\{\pm f(x), 0\}$, then $\|f\| = \int (f^+ - f^-) = \int_{S^+} f^+ + \int_{S^-} f^-$, where $S^\pm$ is the support of $f^\pm$. Now use properties (ii), (iii) and (1) with $A = S^\pm$. A consequence of this is that the eigenvalues of $\mathcal{P}^t$ lie on the unit circle.

2.2 Operator semigroups, generators

The transfer operator also inherits some (semi)group properties of the flow Φ^t.

Definition 2.1. Let $(X, \| \cdot \|)$ be a Banach space. A one parameter family $\{\mathcal{T}^t\}_{t \geq 0}$ of bounded linear operators $\mathcal{T}^t : X \to X$ is called a *semigroup* on X, if

  (a) $\mathcal{T}^0 = I$ (I denoting the identity on X),

  (b) $\mathcal{T}^{t+s} = \mathcal{T}^t \mathcal{T}^s$ for all $t, s \geq 0$.

Further, if $\|\mathcal{T}^t\| \leq 1$, the family is called a semigroup of contractions.
    If $\lim_{t \to 0} \|\mathcal{T}^t f - f\| = 0$ for every $f \in X$, then $\mathcal{T}^t$ is a *continuous semigroup* (C_0 *semigroup*).

    The transfer operator $\mathcal{P}^t$ is a C_0 semigroup of contractions on L^1, see [20] for a proof (in particular Remark 7.6.2 for the continuity).

---

[1] μ is absolutely continuous (with respect to m, if not specified further), if there is an $0 \leq f \in L^1(m)$ such that for all m-measurable $B \subset M$ we have $\mu(B) = \int_B f \, dm$. The function f is called the density of μ.

Definition 2.2. For a semigroup $\mathcal{T}^t$ we define the operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \to X$ by

$$\mathcal{A}f = \lim_{t \to 0} \frac{\mathcal{T}^t f - f}{t}, \qquad f \in \mathcal{D}(\mathcal{A}),$$

with $\mathcal{D}(\mathcal{A}) \subset X$ being the linear subspace of X where the above limit exists. The operator $\mathcal{A}$ is called the *infinitesimal generator* of the semigroup.

For $\mathcal{P}^t$, the infinitesimal generator turns out to be (provided the F_i are continuously differentiable)

$$\mathcal{A}_{PF} f = -\operatorname{div}(fF), \tag{2}$$

see [20]. The following result (see eg. Theorem 2.2.4 [23]) shows the connection between the eigenvalues of the semigroup operators and their infinitesimal generator:

Theorem 2.3 (Spectral mapping theorem)**.** *Let $\mathcal{T}^t$ be a C_0 semigroup and let $\mathcal{A}$ be its infinitesimal generator. Then*

$$e^{t\sigma(\mathcal{A})} \subset \sigma\left(\mathcal{T}^t\right) \subset e^{t\sigma(\mathcal{A})} \cup \{0\},$$

where $\sigma(\cdot)$ denotes the point spectrum of the operator. The corresponding eigenvectors are identical.

This has important consequences for invariant densities:

Corollary 2.4. *The function f is a invariant density of $\mathcal{P}^t$ for all $t \geq 0$ if and only if $\mathcal{A}_{PF} f = 0$.*

According to the discussion at the end of Section 2.1, we have

Corollary 2.5. *The eigenvalues of $\mathcal{A}_{PF}$ lie on the imaginary axis.*

3 The infinitesimal generator, almost-invariant sets, and escape rates

In this section we discuss almost-invariant sets and escape rates via a spectral analysis of the infinitesimal generator. As remarked near the end of Section 2.1, the L^1 spectrum of $\mathcal{P}^t$ lies on the unit circle, and lacks a spectral gap. Applying the spectral mapping theorem, we see that the spectrum of $\mathcal{A}$ must be pure imaginary. In the following discussion, to prove formal results, we will add a small random perturbation to $\mathcal{P}^t$. Later, in the numerics section, we will see that our numerical methods introduce a numerical diffusion that plays the role of creating a spectral gap.

3.1 Stochastic perturbations

In many real world situations, a deterministic model of some physical system is not appropriate. Rather, one should account for the fact that certain external perturbations are present which might be unknown or for which a detailed model would be overly complicated. Often, it is appropriate to account for these influences by incorporating a *small random perturbation* into the description. From a theoretical point of view, this even facilitates the analysis of the system: Under certain assumptions, the transfer operator becomes compact.

We therefore now leave the deterministic setting and assume that our dynamical system described by the ordinary differential equation $dx/dt = F(x)$ is slightly stochastically perturbed by *white noise*. An exact mathematical treatment of this topic would require tools which are beyond the scope of this work (for an exact derivation see [20], chapter 11). We are primarily interested in the time evolution of probability densities. The following material highlights the relevant formal statements and attempts to point out the intuition behind them. Instead of an ordinary differential equation, we now deal with a stochastic differential equation (SDE)

$$\frac{dX}{dt} = F(X) + \varepsilon \frac{dW}{dt}, \tag{3}$$

with $\varepsilon > 0$ and W being a d-dimensional Wiener process. The solutions are time-dependent random variables $X(t)$ with values in $\mathbb{R}^d$. Just as in the deterministic case, we look at the evolution of density functions; now the density functions represent the distribution of random outcomes of the variables $X(t)$. The density functions f satisfy

$$\mathrm{Prob}(X(t) \in A) = \int_A f(x, t) \, \mathrm{d}x.$$

Assuming the existence of such a density $f(x, t)$, where $f(\cdot, 0) = f_0$ is given, we may again define the transfer operator as $\mathcal{P}_\varepsilon^t f_0 := f(t, \cdot)$. If the vector field F is smooth enough we have following characterization of the density function, the *Fokker-Planck* or *Kolmogorov forward equation*, cf. [20], 11.6:

$$\frac{\partial f}{\partial t} = \frac{\varepsilon^2}{2} \Delta f - \mathrm{div}(fF) =: \mathcal{A}_\varepsilon f. \tag{4}$$

Proposition 3.1 ([1, 22]). *The operator $\mathcal{A}_\varepsilon$ (with Neumann boundary conditions[2]) is the infinitesimal generator of C_0 semigroups $\mathcal{P}_{\varepsilon,1}^t$ on L^1, $\mathcal{P}_{\varepsilon,2}^t$ on L^2, and $\mathcal{P}_{\varepsilon,0}^t$ on C^0.*

We note that the operator $\mathcal{P}_{\varepsilon,1}^t$ is identical to the transfer operator $\mathcal{P}_\varepsilon^t$.

Invariant densities Again, we are particularly interested in distributions (densities) which do not change under the evolution. Those are again given by fixed points of $\mathcal{P}_{\varepsilon,1}^t$ or alternatively, by functions in the null space of $\mathcal{A}_\varepsilon$. One can show [30] that $\mathcal{P}_{\varepsilon,1}^t$ is compact and thus the null space of $\mathcal{A}_\varepsilon$ is finite dimensional. Furthermore, due to the white noise, the support of the stochastic transition function associated to $\mathcal{P}_{\varepsilon,1}^t$ is unbounded, and the null space of $\mathcal{A}_\varepsilon$ is one-dimensional, i.e. there is a *unique* invariant density (see again [30], Theorem 1), characterizing the long term dynamical behavior of (3).

3.2   Almost-invariant sets

We call a set $A \subset M$ *almost-invariant* with respect to a (not necessarily invariant) probability measure ν (cf. [15, 14]), if

$$\rho_\nu^t(A) := \frac{\nu\left(\Phi^{-t}(A) \cap A\right)}{\nu(A)} \approx 1 \tag{5}$$

for modest times t. The analogous expression for Φ perturbed by a small random perturbation is

$$\rho_{\nu,\varepsilon}^t(A) := \frac{\mathrm{Prob}_\nu(X(0) \in A, X(t) \in A)}{\mathrm{Prob}_\nu(X(0) \in A)} \approx 1 \tag{6}$$

for modest times t.

We can alternatively characterize this property of a set using an infinitesimal representation of $\mathcal{P}_\varepsilon^t$. To this end let $f \in L^1$ and consider a measurable set $A \subset M$. We define the functional $\mathcal{A}_{\varepsilon,A} : L^1 \supset \mathcal{D}_A(\mathcal{A}_\varepsilon) \to \mathbb{R}$ by

$$\mathcal{A}_{\varepsilon,A} f := \lim_{t \to 0} \int_A \frac{\mathcal{P}_\varepsilon^t f - f}{t} \, dm, \qquad f \in \mathcal{D}_A(\mathcal{A}_\varepsilon), \tag{7}$$

where $\mathcal{D}_A(\mathcal{A}_\varepsilon)$ is the linear subspace of L^1 where the above limit exists. Let $f_A := f\chi_A / (\int f\chi_A dm)$.

Proposition 3.2. *Let $0 \leq f \in \mathcal{D}_A(\mathcal{A}_\varepsilon)$ be the density of the probability measure ν. Then*

$$\rho_{\nu,\varepsilon}^t(A) = 1 + (\mathcal{A}_{\varepsilon,A} f_A)t + o(t) \tag{8}$$

for measurable A as $t \to 0$.[3]

---

[2]The *dynamical* reason why Neumann boundary conditions are chosen here is discussed later.

[3]For two functions $f, g : \mathbb{R} \to \mathbb{R}$ we say "$f(t) = o(g(t))$ as $t \to 0$", if $\lim_{t \to 0} \frac{f(t)}{g(t)} = 0$.

Proof. We have

$$\frac{\rho^t_{\nu,\varepsilon}(A) - 1}{t} = \frac{\mathrm{Prob}_\nu(X(0) \in A, X(t) \in A) - \nu(A)}{t\,\mathrm{Prob}_\nu(X(0) \in A)}$$

$$= \frac{\int_A \mathcal{P}^t_\varepsilon(f\chi_A)\,dm - \int_A f\,dm}{t\,\nu(A)} = \int_A \frac{\mathcal{P}^t_\varepsilon f_A - f_A}{t}\,dm,$$

where the second equation follows from the fact that $X(0)$ is distributed according to ν (with density f), and $X(0) \in A$; which gives the density $f\chi_A$ for $X(0)$, see Section 3.1. Thus for $t \to 0$ the claim follows. $\square$

Correspondingly, a set A will be almost invariant with respect to the measure with density $f \in L^1$, if $\mathcal{A}_{\varepsilon,A} f_A \approx 0$.

There is a strong connection between almost-invariant sets and the spectrum of the generator. The following theorem illustrates this with a simple heuristic for producing almost-invariant sets.

Theorem 3.3. *Suppose that the generator $\mathcal{A}_\varepsilon$ possesses a real eigenvalue $\lambda < 0$ with corresponding eigenfunction f. Then $\int f\,dm = 0$. Define $A^+ = \{f \geq 0\}$, $A^- = \{f < 0\}$. Let $f_{A^+}, f_{A^-} \in \mathcal{D}_A(\mathcal{A})$. Then*

$$\mathcal{A}_{\varepsilon,A^+}|f_{A^+}| + \mathcal{A}_{\varepsilon,A^-}|f_{A^-}| = \lambda.$$

Proof. By Proposition 5.7 in [8] and the spectral mapping theorem we have that

$$\rho^t_{\nu,\varepsilon}(A^+) + \rho^t_{\nu,\varepsilon}(A^-) = \exp(t\lambda) + 1,$$

where ν is the probability measure with density $|f|$. Using Proposition 3.2 we obtain $1 + \mathcal{A}_{\varepsilon,A^+}|f_{A^+}|t + \mathcal{A}_{\varepsilon,A^-}|f_{A^-}|t = \exp(t\lambda) + o(t)$, i.e.

$$\mathcal{A}_{\varepsilon,A^+}|f_{A^+}| + \mathcal{A}_{\varepsilon,A^-}|f_{A^-}| = \frac{\exp(t\lambda) - 1}{t} + o(1)$$

and for $t \to 0$ we obtain the claim. $\square$

For $\lambda \approx 0$ Theorem 3.3 yields $\mathcal{A}_{\varepsilon,A^\pm}|f_{A^\pm}| \approx 0$, which means that the sets A^+ and A^- will be almost-invariant with respect to the probability measure with density $|f|$. Other techniques for extracting sets A^+, A^- from the eigenfunction f may be found in [15, 14]. The papers [15, 14] also discuss almost-invariance with respect to physical *invariant* probability measures, a property that is particularly meaningful when studying typical dynamical behavior. In all cases, the basis for these methods are eigenfunctions of $\mathcal{A}_\varepsilon$ corresponding to (real) eigenvalues close to 0.

If $\mathcal{A}_\varepsilon$ has a complex eigenvalue with real part close to zero, then the corresponding complex eigenfunction may also be used to construct almost-invariant sets.

Lemma 3.4. *Let $\mathcal{A}_\varepsilon f = \lambda_A f$ with $\lambda_A \in \mathbb{C}$ and let f_{re} and f_{im} denote the real and imaginary part of f, respectively. Let $t > 0$ be such that $e^{t\lambda_A} = \lambda_{PF} \in \mathbb{R}$. Then $\mathcal{P}^t_\varepsilon f_{re} = \lambda_{PF} f_{re}$ and $\mathcal{P}^t_\varepsilon f_{im} = \lambda_{PF} f_{im}$.*

Proof. From the proof of Theorem 2.2.4 [23] we have $\mathcal{P}^t_\varepsilon f = \lambda_{PF} f$. Note, that $\mathcal{P}^t_\varepsilon : L^1(M,\mathbb{R}) \circlearrowleft$. By linearity we have $\mathcal{P}^t_\varepsilon f = \mathcal{P}^t_\varepsilon f_{re} + i\,\mathcal{P}^t_\varepsilon f_{im}$. Thus,

$$\underbrace{\lambda_{PF} f_{re}}_{\in \mathbb{R}} + i\,\underbrace{\lambda_{PF} f_{im}}_{\in \mathbb{R}} = \underbrace{\mathcal{P}^t_\varepsilon f_{re}}_{\in \mathbb{R}} + i\,\underbrace{\mathcal{P}^t_\varepsilon f_{im}}_{\in \mathbb{R}}.$$

The claim follows immediately. $\square$

Hence, if for a $t > 0$ we have $1 \approx e^{t\lambda_A} \in \mathbb{R}$, then the real (and imaginary) part of the corresponding eigenfunction yields a decomposition of the phase space into almost invariant sets in the sense of Theorem 3.3.

3.3 Escape rates

We have seen in the previous section that there is a connection between eigenvalues of $\mathcal{A}_\varepsilon$ close to zero and almost-invariant sets. There is also a strong connection between the *escape rate* of the sets A^+, A^- and the corresponding eigenvalue of $\mathcal{A}_\varepsilon$.

In the deterministic setting ($\varepsilon = 0$) the upper (resp. lower) escape rate of a set A under the time-t map of the flow Φ^t is defined as:

$$\bar{E}(A) \quad = \quad -\liminf_{k\to\infty} \frac{1}{k}\log m(A \cap \Phi^{-t}A \cap \cdots \cap \Phi^{-kt}A) \tag{9}$$

$$\underline{E}(A) \quad = \quad -\limsup_{k\to\infty} \frac{1}{k}\log m(A \cap \Phi^{-t}A \cap \cdots \cap \Phi^{-kt}A) \tag{10}$$

If both limits are equal, we call $E(A) = \lim_{k\to\infty} \frac{1}{k}\log m(A \cap \Phi^{-t}A \cap \cdots \cap \Phi^{-kt}A)$ the escape rate from A. The escape rate is the asymptotic exponential rate of loss of m-mass from the set A. One might expect that almost-invariant sets have low escape rate and vice-versa, but simple counterexamples in [17] show that this is not the case. This is because the notion of almost-invariance as defined above is a finite-time property, while escape rate is an asymptotic quantity.

The $\varepsilon > 0$ version of (9) is

$$\bar{E}_\varepsilon(A) \quad := \quad -\liminf_{k\to\infty} \frac{1}{k}\log \mathrm{Prob}_m(X(0) \in A, X(t) \in A, \ldots, X(kt) \in A) \tag{11}$$

$$= \quad -\liminf_{k\to\infty} \frac{1}{k}\log \int_A (\mathcal{P}^t_{\varepsilon,A})^k(1)\,dm. \tag{12}$$

where $\mathcal{P}^t_{\varepsilon,A}(f) := \mathcal{P}^t_\varepsilon(f\chi_A)$ is the standard restriction of the operator $\mathcal{P}^t_\varepsilon$ to A [24]. The following theorem provides a recipe for constructing sets with low escape rate from eigenfunctions of $\mathcal{A}_\varepsilon$ with real eigenvalues.

Theorem 3.5. *Suppose that $\mathcal{A}_\varepsilon f = \lambda f$ for some $\lambda < 0$ and $f \in L^\infty(m)$. Then*

$$\bar{E}_\varepsilon(A^+) \leq -t\lambda \text{ and } \bar{E}_\varepsilon(A^-) \leq -t\lambda,$$

where $A^+ = \{f \geq 0\}$ and $A^- = \{f < 0\}$.

Proof. Let f be scaled such that $\int |f|\,dm = 2$, and define the (signed) measure ν as $\nu(A) = \int_A f\,dm$. Then $\nu^+ = \nu|_{A^+}$ and $\nu^- = -\nu|_{A^-}$ are both probability measures. For an event $\mathcal{E}$ define $\mathrm{Prob}_\nu(\mathcal{E}) := \mathrm{Prob}_{\nu^+}(\mathcal{E}) - \mathrm{Prob}_{\nu^-}(\mathcal{E})$. Since f is an eigenfunction of $\mathcal{P}^t_\varepsilon$ with eigenvalue $e^{\lambda t}$, we have

$$
\begin{aligned}
e^{\lambda k t} = e^{\lambda k t}\int_{A^+} f\,dm \quad &= \quad \int_{A^+} \mathcal{P}^{kt}_\varepsilon f\,dm = \mathrm{Prob}_\nu\big(X(kt) \in A^+\big) \\
&= \quad \mathrm{Prob}_\nu\big(X(\ell t) \in A^+, \ell = 0, \ldots, k\big) \\
&\quad + \sum_{n=0}^{k-1} \underbrace{\mathrm{Prob}_\nu\big(X(nt) \in A^-, X(\ell t) \in A^+, \ell = n+1, \ldots, k\big)}_{=:p_n}.
\end{aligned}
$$

It follows for the summands p_n in the above sum:

$$
\begin{aligned}
p_n \quad &= \quad \mathrm{Prob}_{\mathcal{P}^{nt}_\varepsilon \nu}\big(X(0) \in A^-, X(\ell t) \in A^+, \ell = 1, \ldots, k-n\big) \\
&= \quad e^{\lambda n t}\mathrm{Prob}_\nu\big(X(0) \in A^-, X(\ell t) \in A^+, \ell = 1, \ldots, k-n\big) \\
&= \quad -e^{\lambda n t}\mathrm{Prob}_{\nu^-}\big(X(\ell t) \in A^+, \ell = 1, \ldots, k-n\big),
\end{aligned}
$$

where in the first equation the action of the transfer operator $\mathcal{P}^t_\varepsilon$ on the measure ν is defined through its action on the density f of ν. The second equation follows from ν being an eigenmeasure. Clearly, p_n is non-positive, and thus

$$\mathrm{Prob}_\nu\big(X(\ell t) \in A^+, \ell = 0, \ldots, k\big) \geq e^{\lambda k t}$$

holds. The rest of the proof follows the lines of the one of Theorem 2.4 in [17]. $\square$

7

The recipe of Theorem 3.5 is in fact the same as in Theorem 3.3. The difference is the measure used: in Theorem 3.3, almost-invariance is computed with respect to the particular measure ν with density $|f|$ where f is the eigenfunction in question, whereas in Theorem 3.5 escape is computed with respect to Lebesgue measure.

Remark 3.6. A more natural notion of escape rates in the time-continuous case would be the following:

$$\bar{E}_\varepsilon(A) := -\liminf_{t \to \infty} \frac{1}{t} \text{Prob}_m\big(X(s) \in A, s \in [0, t)\big).$$

However, for this definition it is more complicated to obtain similar results to the one in Theorem 3.5, and a fuller discussion will appear elsewhere. One can view (11) with $t = 1$ as the "time sampled version" of this continuous-time definition.

4 Numerical approximation

Having seen that certain eigenpairs of the transfer operator (resp. infinitesimal generator) carry the information we are seeking, we describe in the sequel our proposed approximation of these eigenpairs. To this end, we define finite dimensional approximation spaces and consider the eigenvalue problem projected onto these spaces. Throughout this section we assume that the underlying deterministic vector field F is smooth.

4.1 Ulam's method

We describe here the "standard" Ulam approach; see the surveys [13, 6] for more details. We partition M into d-dimensional connected, positive volume subsets $\{B_1, \ldots, B_n\}$. Typically, each B_i will be a hyperrectangle or simplex to simplify computations. As an approximation space we consider the space $\Delta_n = \text{sp}\{\chi_{B_1}, \ldots, \chi_{B_n}\}$ of functions which are piecewise constant on the cells of the partition. Let $\pi_n : L^1 \to \Delta_n$, $\pi_n f = \sum_{i=1}^n \frac{1}{m(B_i)} \int_{B_i} f \ dm \ \chi_{B_i}$, be the L^2-orthogonal projection onto Δ_n. We let $\mathcal{P}_n^t : \Delta_n \to \Delta_n$, $\mathcal{P}_n^t := \pi_n \mathcal{P}^t$, be the approximate Frobenius-Perron operator. Note that $\mathcal{P}_n^t \chi_{B_i} = \pi_n \mathcal{P}^t \chi_{B_i} = \sum_{j=1}^n \frac{1}{m(B_j)} \int_{B_j} \mathcal{P}^t \chi_{B_i} \ dm \ \chi_{B_j}$, i.e. the matrix representation $P_n^t \in \mathbb{R}^{n \times n}$ of $\mathcal{P}_n^t$ with respect to the basis $\chi_{B_1}, \ldots, \chi_{B_n}$ and multiplication on the left is

$$(P_n^t)_{ij} = \frac{1}{m(B_j)} \int_{B_j} \mathcal{P}^t \chi_{B_i} \ dm = \frac{m(B_i \cap \Phi^{-t} B_j)}{m(B_j)}.$$

This matrix is easily constructed numerically using eg. GAIO [6].

In the stochastic setting, letting $\mathcal{P}_{\varepsilon,n}^t : \Delta_n \to \Delta_n$, $\mathcal{P}_{\varepsilon,n}^t := \pi_n \mathcal{P}_\varepsilon^t$, one obtains

$$(P_{\varepsilon,n}^t)_{ij} = \frac{1}{m(B_j)} \int_{B_j} \mathcal{P}_\varepsilon^t \chi_{B_i} \ dm$$

as above. In principle one can use Monte-Carlo integration to compute these entries, however, this is not particularly efficient.

4.2 Ulam's method for the generator

We partition M as in the standard Ulam's method. We will see that the numerical scheme itself introduces some diffusion and we therefore consider the deterministic generator $\mathcal{A}$ (i.e. $\varepsilon = 0$).

We wish to construct an operator $\mathcal{A}_n : \Delta_n \to \Delta_n$ that is close in some sense to the operator $\mathcal{A}$. Motivated by Ulam's method, one would like to form $\pi_n \mathcal{A}$, which unfortunately does not exist, because $\Delta_n \not\subseteq \mathcal{D}(\mathcal{A})$, cf. (2). Instead of differentiating w.r.t. time and then doing the projection, we swap the order of these operations. Let us build the Ulam approximation $\mathcal{P}_n^t$ first, which will

not be a semigroup any more, but for fixed t it approximates $\mathcal{P}^t$. Taking the time derivative, our candidate approximate operator is

$$\mathcal{A}_n f := \lim_{t \to 0} \left(\frac{\pi_n \mathcal{P}^t \pi_n f - \pi_n f}{t} \right). \tag{13}$$

The following lemma emphasizes the intuition behind this definition: if $\mathcal{P}_n^t$ is a sufficiently good approximation (for small t) of the Markov jump process generated by the dynamics on the sets B_i, then $\mathcal{A}_n$ will be the generator of this process. To be exact, the following is the case:

Lemma 4.1. *The matrix representation of* $\mathcal{A}_n : \Delta_n \circlearrowleft$ *with respect to the basis* $\chi_1, \dots, \chi_n$ *under multiplication on the left is*

$$(A_n)_{ij} = \begin{cases} \displaystyle\lim_{t \to 0} \frac{m(B_i \cap \Phi^{-t} B_j)}{t \cdot m(B_j)}, & i \neq j; \\ \displaystyle\lim_{t \to 0} \frac{m(B_i \cap \Phi^{-t} B_i) - m(B_i)}{t \cdot m(B_i)}, & otherwise. \end{cases} \tag{14}$$

Proof. We consider the action of $\mathcal{P}^t$ on χ_{B_i}.

$$
\begin{aligned}
\lim_{t \to 0} \pi_n \frac{\mathcal{P}^t \chi_{B_i} - \chi_{B_i}}{t} &= \lim_{t \to 0} \sum_{j=1}^n \frac{1}{m(B_j)} \left(\int_{B_j} \frac{\mathcal{P}^t \chi_{B_i} - \chi_{B_i}}{t} \, dm \right) \chi_{B_j} \\
&= \lim_{t \to 0} \sum_{j \neq i} \frac{1}{m(B_j)} \left(\int_{B_j} \frac{\mathcal{P}^t \chi_{B_i}}{t} \, dm \right) \chi_{B_j} \\
&\quad + \lim_{t \to 0} \frac{1}{m(B_i)} \left(\int_{B_i} \frac{\mathcal{P}^t \chi_{B_i} - \chi_{B_i}}{t} \, dm \right) \chi_{B_i} \\
&= \lim_{t \to 0} \sum_{j \neq i} \frac{1}{m(B_j)} \left(\int_{\Phi^{-t} B_j} \frac{\chi_{B_i}}{t} \, dm \right) \chi_{B_j} \\
&\quad + \lim_{t \to 0} \frac{1}{m(B_i)} \left(\int_{\Phi^{-t} B_i} \frac{\chi_{B_i}}{t} \, dm - \int_{B_i} \frac{\chi_{B_i}}{t} \, dm \right) \chi_{B_i} \\
&= \sum_{j \neq i} \lim_{t \to 0} \frac{m(B_i \cap \Phi^{-t}(B_j))}{t \cdot m(B_j)} \chi_{B_j} + \lim_{t \to 0} \frac{m(B_i \cap \Phi^{-t} B_i) - m(B_i)}{t \cdot m(B_i)} \chi_{B_i}
\end{aligned}
$$

Thus under left multiplication we obtain (14). $\qquad\square$

Remark 4.2. Lemma 4.1 states that $(A_n)_{ij}$ is the *outflow rate of mass* from B_i into B_j.

Let $\mathcal{R}_n^t := \exp(t\mathcal{A}_n) = I - \pi_n + \exp(t\mathcal{A}_n \mid_{V_n})\pi_n$ denote the semigroup generated by $\mathcal{A}_n$. Then $\mathcal{R}_n^t$ and $\mathcal{P}^t$ are near to each other in the following sense, cf. [18]:

Proposition 4.3. *As* $t \to 0$

$$\mathcal{R}_n^t f - \pi_n \mathcal{P}^t f = \mathcal{O}(t^2) \tag{15}$$

for all $f \in \Delta_n$.[4]

The following lemma allows us to construct $\mathcal{A}_n$ without the computation of the flow Φ^t.

Lemma 4.4. *For* $i \neq j$, *define* $\mathbf{n}_{ij}$ *to be the the unit normal vector pointing out of* B_i *into* B_j *if* $B_i \cap B_j$ *is a* $d-1$-*dimensional face, and the zero vector otherwise. The matrix representation of* $\mathcal{A}_n : \Delta_n \circlearrowleft$ *with respect to the basis* $\chi_1, \dots, \chi_n$ *under multiplication on the left is*

$$(A_n)_{ij} = \begin{cases} \displaystyle\frac{1}{m(B_j)} \int_{B_i \cap B_j} \max\{F(x) \cdot \mathbf{n}_{ij}, 0\} \, dm_{d-1}(x), & i \neq j; \\ \displaystyle-\sum_{j \neq i} \frac{m(B_j)}{m(B_i)} (A_n)_{ij}, & otherwise. \end{cases} \tag{16}$$

---

[4]For two functions $f, g : \mathbb{R} \to \mathbb{R}$ we say "$f(t) = \mathcal{O}(g(t))$ as $t \to 0$", if $\limsup_{t \to 0} \frac{|f(x)|}{|g(x)|} < \infty$.

Proof. From (14) we have for $i \neq j$ that $A_{n,ij} = \lim_{t \to 0} \frac{m(B_i \cap \Phi^{-t}B_j)}{t \cdot m(B_j)}$. Denoting $M_{ij}(t) = m(B_i \cap \Phi^{-t}B_j)$ we have that $A_{n,ij} = M'_{ij}(0)/m(B_j)$ where the prime denotes differentiation with respect to t. The quantity $M'_{ij}(0)$ is simply the rate of flux out of B_i through the face $B_i \cap B_j$ into B_j and so $M'_{ij}(0) = \int_{B_i \cap B_j} \max\{F(x) \cdot \mathbf{n}_{ij}, 0\} dm_{d-1}(x)$.

For the diagonal elements A_n we have $A_{n,ii} = \lim_{t \to 0} \frac{m(B_i \cap \Phi^{-t}B_i) - m(B_i)}{t \cdot m(B_i)}$. Note that $m(B_i) - m(B_i \cap \Phi^{-t}B_i) = m(B_i \setminus \Phi^{-t}B_i)$. Clearly $B_i \setminus \Phi^{-t}B_i = B_i \cap \bigcup_{j \neq i} \Phi^{-t}B_j = \bigcup_{j \neq i} B_i \cap \Phi^{-t}B_j$ modulo sets of Lebesgue measure zero. Thus, $m(B_i) - m(B_i \cap \Phi^{-t}B_i) = \sum_{j \neq i} m(B_i \cap \Phi^{-t}B_j)$. Now, by (14), $A_{n,ii} = -\lim_{t \to 0} \frac{\sum_{j \neq i} m(B_i \cap \Phi^{-t}B_j)}{m(B_i)} = -\sum_{j \neq i} \frac{m(B_j)}{m(B_i)} A_{n,ij}$. $\square$

In one dimension, (16) has a particularly simple form.

Corollary 4.5. *Let $M = \mathbb{T}^1$. Assume (without loss[5]) that $F(x) \geq 0$. Let $0 = x_0 < x_1 < \cdots < x_n = 1$ and consider the partition $\{B_1, \ldots, B_n\}$ with $B_i = [x_{i-1}, x_i]$. Then*

$$(A_n)_{ij} = \begin{cases} F(x_i)/m(B_j), & j = i+1; \\ -F(x_i)/m(B_i), & j = i; \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Remark 4.6 (Connections with the upwind scheme)**.** Clearly, $\mathcal{A}_n$ is the spatial discretization from the so-called upwind scheme in finite volume methods; cf. [21]. The scheme is known to be stable. Stability of finite volume schemes is often related to "numerical diffusion" in them.

We demonstrate numerical diffusion on a simple example, for further details we refer to [21] Section 8.6.1. Set $M = \mathbb{T}^1$ and $F(x) = \hat{F} > 0 \; \forall x \in M$. Then $\mathcal{A}f = -\hat{F}f'$ and $\mathcal{A}_n$ is the backward difference scheme. Consider $f_n := \pi_n f$ as a vector of values. For sufficiently smooth f we have

$$(\mathcal{A}_n f)_i = n\hat{F}(f_{n,i-1} - f_{n,i}) = (\pi_n \mathcal{A}f)_i + \mathcal{O}(n^{-1}),$$

but

$$\begin{aligned} (\mathcal{A}_n f)_i = n\hat{F}(f_{n,i-1} - f_{n,i}) &= n\hat{F}\left(\frac{f_{n,i-1} - f_{n,i+1}}{2} + \frac{f_{n,i-1} - 2f_{n,i} + f_{n,i+1}}{2}\right) \\ &= \hat{F}\frac{f_{n,i-1} - f_{n,i+1}}{2n^{-1}} + \frac{\hat{F}}{2n}\frac{f_{n,i-1} - 2f_{n,i} + f_{n,i+1}}{n^{-2}} \\ &= \left(\pi_n\left(\mathcal{A}f + \frac{\hat{F}}{2n}\Delta f\right)\right)_i + \mathcal{O}(n^{-2}), \end{aligned}$$

hence $\mathcal{A}_n f$ is a better approximation of $\mathcal{A}_\varepsilon f$, with $\frac{\varepsilon^2}{2} = \frac{\hat{F}}{2n}$, and $\mathcal{A}_\varepsilon$ is the infinitesimal generator associated with the SDE $\frac{dx}{dt} = F(x) + \varepsilon\frac{dW}{dt}$ than of $\mathcal{A}f$. That is why one expects quantities computed by $\mathcal{A}_n$ to reflect the actual behavior of $\mathcal{A}_\varepsilon$. Since the above equations contain only *local* estimates, for a general nonconstant F is the diffusion also spatially varying; i.e. $\frac{\varepsilon(x)^2}{2} = \frac{F(x)}{2n}$. Note, that this notion of numerical diffusion is easily extended to multiple space dimensions.

The definition (13) allows us to interpret the numerical diffusion in yet another way. We showed in Proposition 4.3 that P_n^t is the transition matrix of a Markov process near the Markov jump process generated by A_n for small $t > 0$. The discretized FPO P_n^t can be related to a non-deterministic dynamical system, which, after mapping the initial point, adds some uncertainty to produce a uniform distribution of the image point in the box where it landed; see [12]. Thus, this uncertainty resulting from the numerical discretization, equivalent to the numerical diffusion in the upwind scheme, can be viewed as the reason for robust behavior — stability.

Finally, we show that our constructions (16) and (17) always provide a solution to the system $\mathcal{A}_n f = 0$ for some $f \in \Delta_n$.

---

[5]If $F \not\geq 0$ and $F \not\leq 0$, we have one or more stable fixed points, and every trajectory converges to one of them. Hence, there is no interesting statistical behavior to analyze.

Lemma 4.7. *There exists a nonnegative, nonzero $f \in \Delta_n$ so that $\mathcal{A}_n f = 0$.*

Proof. Let $D_{n,ij} = m(B_i)\delta_{ij}$ and note that $Q_n := D_n^{-1} A_n D_n$ satisfies

$$(Q_n)_{ij} = \begin{cases} \frac{m(B_j)}{m(B_i)} A_{ij}, & i \neq j; \\ -\sum_{j \neq i} \frac{m(B_j)}{m(B_i)}(A_n)_{ij}, & \text{otherwise.} \end{cases} \tag{18}$$

Note that all row sums of Q_n equal zero. Let $c = \sum_{j \neq i} Q_{n,ij}$. The matrix $\hat{Q}_n := Q_n + cI$ is nonnegative with all row sums equal to c. By the Perron-Frobenius theorem [3], the largest eigenvalue of $\hat{Q}_n$ is c (of multiplicity[6] possibly greater than 1) and one of the corresponding left eigenvectors pair u_n is nonnegative. Clearly u_n is a left eigenvector of Q_n corresponding to the eigenvalue 0 and thus $D_n u_n$ is a nonnegative left eigenvector corresponding to 0 for A_n. $\qquad \square$

4.3   Spectral collocation for the generator

The eigenfunctions of $\mathcal{A}_\varepsilon$ are smooth. This motivates the use of smooth approximation functions, e.g. polynomials. We here outline the general principles of spectral collocation methods; for a more thorough presentation we refer to [5], [4] or [27].

Choose a family of approximation spaces $\{V_n\}_{n \in \mathbb{N}}$, such that $V_n \subset C^\infty(M)$ for all n. Depending on the type of the phase space, we use two different approximation spaces. We introduce them in one dimension; the multidimensional ones can then be constructed by tensor products. In both cases, the approximation space comes with an associated set of collocation nodes:

- *Periodic domain/uniform grid.* We have $M = \mathbb{T}^1$ and restrict ourselves to odd values of n. Then the basis we choose for V_n is

$$\left\{ e^{ikx} \right\}_{-n/2 - 1 \leq k \leq n/2}.$$

  The associated collocation nodes are the uniform grid $\{0, 1/n, \ldots, (n-1)/n\}$.

- *Standard domain/Chebyshev grid.* Here, $M = [-1, 1]$. The space V_n is spanned by the monomials of order 0 to n. We use Chebyshev polynomials as basis functions:

$$\left\{ \cos\left(k \arccos\left(x\right)\right) \right\}_{0 \leq k \leq n},$$

  together with the Chebyshev grid $\{-\cos(2\pi j/n), j = 0, \ldots, n\}$, as collocation nodes.

Let $f \in V_n$ and $\mathcal{I}_n : C^\infty \to V_n$ be the interpolation operator for the given collocation nodes. We define the approximate generator by

$$\mathcal{A}_{\varepsilon,n} f := \mathcal{I}_n \mathcal{A}_\varepsilon f.$$

For both cases we have following:

Theorem 4.8 (Spectral accuracy, [5])**.** *For $f \in C^\infty(M)$ let f_n be the best approximation of f in V_n w.r.t. the supremum norm $\| \cdot \|_\infty$. Then for each $k \in \mathbb{N}$ there is a $c_k > 0$ such that*

$$\|f - f_n\|_\infty \leq c_k \, n^{-k} \qquad \text{for all } n \in \mathbb{N}. \tag{19}$$

Convergence then follows from standard results on the analysis of Galerkin methods for elliptic differential operators and spectral approximation (cf. also [18]):

Theorem 4.9. *Let M be a compact tensor product domain with smooth boundary. Let F be a smooth vector field on M and $\mathcal{A}_\varepsilon$, $\mathcal{A}_{\varepsilon,n}$ defined as above, with Neumann (resp. periodic) boundary conditions. Then the eigenvalues and eigenfunctions of $\mathcal{A}_\varepsilon$ are approximated by the corresponding eigenvalues and eigenfunctions of $\mathcal{A}_{\varepsilon,n}$ with spectral accuracy.*

---

[6]If $\hat{Q}_n$ is aperiodic (there exists k such that $\hat{Q}_n^k > 0$) then the eigenvalue c is simple and the corresponding eigenvector is positive (see [3]).

Why Neumann boundary conditions? The objects that we are approximating are densities, so there should be no "loss of mass" as time goes on. In a closed physical system, the flow normal to the boundary is zero. Hence, there is no loss of mass through advection. The physical meaning of diffusion is a flow of mass in the direction opposite to the gradient and proportional to its magnitude. Thus, no loss of mass via diffusion translates into Neumann boundary conditions for the densities: their normal gradients have to be zero at the boundary. Equivalently, one may write (4) as a continuity equation,

$$\partial_t f = \text{div}(I),$$

with $I := \frac{\varepsilon^2}{2} \nabla f - fF$ being the probability flow or probability current. The condition $I|_{\partial M} = 0$ leads precisely to Neumann boundary conditions.

4.4 Algorithms

Algorithm 1 (Ulam's method)**.**

1. Partition M into connected sets $\{B_1, \ldots, B_n\}$ of positive volume. Typically, each B_i will be a hyperrectangle or simplex.

2. Choose $t > 0$ and compute the matrix P_n^t or $P_{\varepsilon,n}^t$ as described in Section 4.1.

3. Estimates of invariant densities for Φ^t are given by fixed points of P_n^t (or $P_{\varepsilon,n}^t$): Let $vP_n^t = v$. Then $f := \sum_{i=1}^n v_i \chi_{B_i}$ satisfies $\mathcal{P}_n^t f = f$.

4. Similarly, eigenvectors of P_n^t corresponding to real eigenvalues $\lambda \approx 1$ provide information about almost-invariant sets (cf. Theorem 3.3) and sets with low escape rates (cf. Theorem 3.5).

Note that P_n^t typically is a sparse matrix and the number of nonzero entries per column will be determined by Lipschitz constants of Φ^t.

Algorithm 2 (Ulam's method for the generator)**.**

1. Partition M into connected sets $\{B_1, \ldots, B_n\}$ of positive volume. Typically each B_i will be a hyperrectangle or simplex.

2. Compute

$$
(A_n)_{ij} = \begin{cases} \dfrac{1}{m(B_j)} \displaystyle\int_{B_i \cap B_j} \max\{F(x) \cdot \mathbf{n}_{ij}, 0\} \, dm_{d-1}(x), & i \neq j, \\[2ex] -\displaystyle\sum_{j \neq i} \dfrac{m(B_j)}{m(B_i)}(A_n)_{ij}, & \text{otherwise,} \end{cases}
$$

where one uses standard quadrature rules to estimate the integral.

3. Estimates of invariant densities lie in the null space of A_n. Let $vA_n = 0$ (the existence of such a v is guaranteed by Lemma 4.7), then $f := \sum_{i=1}^n v_i \chi_{B_i}$ satisfies $\mathcal{A}_n f = 0$.

4. Similarly, eigenvectors of A_n corresponding to large real eigenvalues $\lambda < 0$ provide information about almost-invariant sets and sets with low escape rates.

Note that the discretized generator A_n is a sparse matrix since $A_{n,ij} = 0$ if B_i and B_j do not share a boundary.

Algorithm 3 (Spectral collocation for the generator)**.**

1. Set up a grid $x_0, \ldots, x_n$, which is the Chebyshev grid $x_j = -\cos(2\pi j/n)$, $j = 0, \ldots, n$, if $M = [-1, 1]$, or the equispaced grid $x_j = j/n$, if $M = \mathbb{T}^1$. (If the state space M is an affine transformation of the above ones, the grid is transformed analogously.) For multidimensional tensor product spaces the grid is the tensor product of the one dimensional grids.

2. Let $\{\ell_0, \ldots, \ell_n\}$ denote the Lagrange basis on the grid, i.e. all ℓ_i are polynomials of maximal degree n with $\ell_i(x_j) = \delta_{ij}$. Denote the interpolation on the grid by $\mathcal{I}_n$, and define the discretization matrix obtained by collocation as

$$(A_{\varepsilon,n})_{ij} = (\mathcal{I}_n \mathcal{A}_\varepsilon \ell_j)(x_i).$$

Note, that the computation of these matrix entries is very simple. If $M = \mathbb{T}^1$, we may switch between the evaluation space and the frequency space simply by the fast Fourier transformation. Multiplication by the vector field F is pointwise multiplication in the evaluation space, computing the derivative is a diagonal scaling in the frequency space. This can be simply extended to $M = [0, 1]$ (cf. [27] Chapter 8), as well as for a multidimensional phase space.

3. Compute the left eigenvector v of $A_{\varepsilon,n}$ at the eigenvalue of smallest magnitude. Then, $\sum_{i=0}^{n} v_i \ell_i$ approximates the invariant density.

4. Similarly, eigenvectors of $A_{\varepsilon,n}$ corresponding to large real eigenvalues $\lambda < 0$ provide information about almost-invariant sets.

4.5 Solving the eigenproblem

Once the matrix approximation of the operator or the generator has been computed, we then have to solve the corresponding eigenvalue problem. In Ulam's method one seeks dominant eigenvalues, and standard Arnoldi type methods easily provide the interesting part of the spectrum (we use the `eigs` function in Matlab).

For the approximate generator, we look for eigenvalues with small modulus. Here, the Arnoldi iteration in `eigs` uses inverse iteration, i.e. repeatedly solves linear systems of the type $A_n u = b$. Since the matrix A_n is singular, this is not a well posed problem and we might expect some difficulties with these iterative methods. A possibility for overcoming them is to solve the eigenvalue problem for $\bar{A}_n := A_n - \sigma I_{n \times n}$, with a suitable shift σ. This merely introduces a shift of the spectrum and the eigenfunctions stay the same. For n big enough[7], the spectrum of $\bar{A}_n$ is going to be well separated from zero. The size n of the linear system may be large, so a direct solution via LU factorization may not be feasible. However, in Ulam's method for the generator, A_n is a sparse matrix, hence iterative methods (e.g. GMRES) can be used.

For spectral collocation, in general, the matrix A_n will not be sparse. Nevertheless, in the case where the domain is periodic, the matrix-vector product $A_n x$ can be computed fast using the FFT.

4.6 Computational complexity

For d dimensional systems in general, (16) shows us that setting up the approximate generator requires the computation of $(d-1)$- dimensional integrals, while to set up the Ulam-matrix, we compute d-dimensional integrals. Of course, the gain is biggest in one dimension (zero dimensional integrals are one-point-evaluations). Equally importantly, (16) does not require integration of the vector field.

For Ulam's method and Ulam's method for the generator, the matrix entries should be computed to an accuracy $\mathcal{O}(n^{-1/d})$; otherwise we cannot expect the approximate eigenfunctions to

---

[7]If A_n is obtained from the Ulam-based approach, the spectrum of A_n, which is the generator of a Markov jump process, is in the left complex half plane. This does not hold in general for the spectral collocation approach, here we rely on the fast convergence of eigenvalues.

exploit the full potential of the approximation space (which is an error of $\mathcal{O}(n^{-1/d})$ for n partition elements in a d dimensional space). Assuming the quadrature rule used to compute the entries to also suffer from the curse of dimension, we expect to need $\mathcal{O}(n)$ and $\mathcal{O}(n^{(d-1)/d})$ function evaluations for each matrix entry in the Ulam-matrix, and the matrix arising from Ulam's method for the generator, respectively. This explains the third column of Table 1 below.

5   Examples

A collection of examples is presented in the following. The examples range over three phase space dimensions, different boundary types, and chaotic and regular dynamics. We find that the generator approaches are typically superior to the standard Ulam approach, producing more accurate results with less computation time. If the vector field is infinitely differentiable, the spectral collocation approach can be extremely accurate. In our final example, we find that for flows on complicated attractors, the standard Ulam method can take advantage of the attracting dynamics to produce better estimates of invariant densities than the generator approaches. In all cases, the operator based methods are far superior to direct trajectory integration for estimating the invariant density. Other structures such as almost-invariant sets can only be identified using the operator approaches.

5.1   A flow on the circle

We start with a one dimensional example, a flow on the unit circle. The vector field is given by

$$F(x) = \sin(4\pi x) + 1.1,$$

$x \in \mathbb{T}^1 = [0,1]$ with periodic boundary conditions, and we wish to compute the invariant density of the system. Recall that an invariant density $f \in L^1(\mathbb{T}^1)$ needs to satisfy $\mathcal{A}f = 0$, where $\mathcal{A}f(x) = -\,\mathrm{d}(fF)(x)/\,\mathrm{d}x$. The unique solution to this equation is $f^*(x) = C/F(x)$, C being a normalizing constant (i.e. such that $\|f\|_1 = 1$). The honours thesis [26] numerically investigated the estimation of the invariant density for this flow, and other two-dimensional flows, by finding approximate eigensolutions of the infinitesimal generator using finite difference and finite element methods. Here we use the three methods discussed in the previous section to approximate f^* and compare their efficiency relative to one another and a histogram of a long simulation:

1. the classical method of Ulam for the Frobenius-Perron operator,

2. Ulam's method for the generator and

3. spectral collocation for the generator.

As we have a periodic domain and the vector field is infinitely smooth, we expect spectral collocation to perform very well. Figure 1 shows the true invariant density (dashed line), together with its approximations by the four methods.

5.1.1   Matlab code

In this 1D case, the three methods can each be realized in a few lines of Matlab. For illustration purposes, we include the code here.

Ulam's method

```
F = @(x) sin(4*pi*x)+1.1;                    % definition of the vector field
n = 32; x = 1/(2*n^2):1/(n^2):1-1/(2*n^2);   % nodes for spatial integration
[t,y] = ode23(@(t,x) F(x),[0 2/n],x');       % time integration of the nodes
I = ceil(max(n*mod(y(end,:),1),1));          % cell indices of image points
J = reshape(ones(n,1)*(1:n),1,n*n);          % construction of transition matrix
P = sparse(I,J,1/n,n,n);
```

14

```
[v,d] = eig(full(P));                          % spectrum of transition matrix
f_n = @(x) abs(v(ceil(max(x*n,1)),1));         % plot approximate inv. density
L1 = quadl(f_n,0,1); fplot(@(x)f_n(x)/L1,[0,1])
```

The error in Ulam's method decreases like $\mathcal{O}(n^{-1})$ for smooth invariant densities [10]. Thus, we need to compute the transition rates between the intervals to an accuracy of $\mathcal{O}(n^{-1})$ (since otherwise we cannot expect the approximate density to have a smaller error). To this end, we use a uniform grid of n sample points in each interval. This leads to $\mathcal{O}(n^2)$ evaluations of the vector field. For the numbers in Figure 1 we only counted each point once, i.e. we neglected the fact that for the time integration we have to perform several time steps per point.
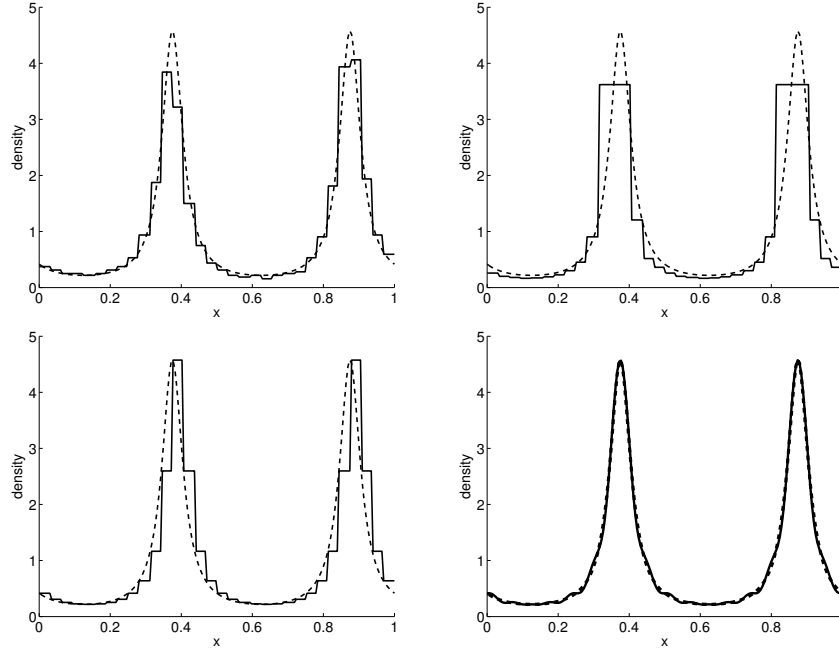

Figure 1: True invariant density (dashed line), approximation by histogramming a long simulation (top left, 1024 iterates), Ulam's method (top right, 32 sample points per partition element, 32 partition elements), Ulam's method for the generator (bottom left, 32 partition elements) and spectral collocation (bottom right, 32 collocation points).

Ulam's method for the generator

```
F = @(x) sin(4*pi*x)+1.1;                          % definition of the vector field
n = 32; Fx = F(0:1/n:1-1/n);                       % evaluation on the boundary nodes
A = n*sparse(1:n,1:n,-Fx) + sparse(1:n,[2:n 1],Fx); % assembling the discrete generator
[v,d] = eig(full(A'));                             % spectrum
[d,I] = sort(diag(d));
f_n = @(x) abs(v(ceil(max(x*n,1)),I(1)));          % plot approximate inv. density
L1 = quadl(f_n,0,1); fplot(@(x)f_n(x)/L1,[0,1])
```

Here, only one evaluation of the vector field per interval is needed as in one dimension the integration on the boundaries reduces to the evaluation of a single boundary point. On a partition with n intervals, this method then yields an accuracy of $\mathcal{O}(n^{-1})$. Note that from Corollary 4.5 it follows that the vector v with $v_i = 1/F(x_i)$ is a left eigenvector of the transition matrix (17) for the generator at the eigenvalue 0. This fact proves pointwise convergence of the invariant density of the discretization towards the real one. Thus Ulam's method for the generator is very accurate and stable; the L^1 error is $\mathcal{O}(1/n)$ as this is the rate at which approximants created from a basis of characteristic functions converge in L^1 to regular functions.

Spectral collocation

```
F = @(x) sin(4*pi*x)+1.1;                % definition of the vector field
n = 32; Fx = F(0:1/n:1-1/n);             % evaluation in collocation nodes
E = ifft(eye(n));                        % basis
FE = fft(diag(Fx)*E);                    % multiplication by vector field
I = [0:n/2-1 -n/2:-1];                   % frequencies
D = (2i*pi)*I';                          % differentiation matrix
A = -D*ones(1,n).*FE;                    % discrete generator in frequency space
[v,lambda] = eig(A);                     % spectrum
f_n = @(x) real(exp(x'*I*2i*pi)*v(:,end)); % approximate density
L1 = quadl(f_n,0,1); fplot(@(x)f_n(x)/L1,[0,1]) % plot
```

Here, the vector field is evaluated once per grid point (cf. the second line of the code). As predicted by Theorem 4.8, the accuracy increases exponentially with n (cf. Figure 2).

5.1.2 Computational efficiency

In Figure 2 (left) we compare the efficiency of the four methods in terms of how the L^1-error of the computed invariant density depends on the number of evaluations of the vector field. The L^1-error was computed using an adaptive Lobatto quadrature as implemented in, e.g., Matlab's `quadl` command. These errors are due to the approximation space, the approximate numerical computation of matrix entries, and due to solving the discretized eigenvalue problem. To illustrate how these effects accumulate, in Figure 2 (right) we compare the errors of three different approximations, for different (uniform) partitions: the approximate invariant densities obtained from the two Ulam type methods, and the projection of the true invariant density. We conclude that the error due to the approximation space dominates the total error.
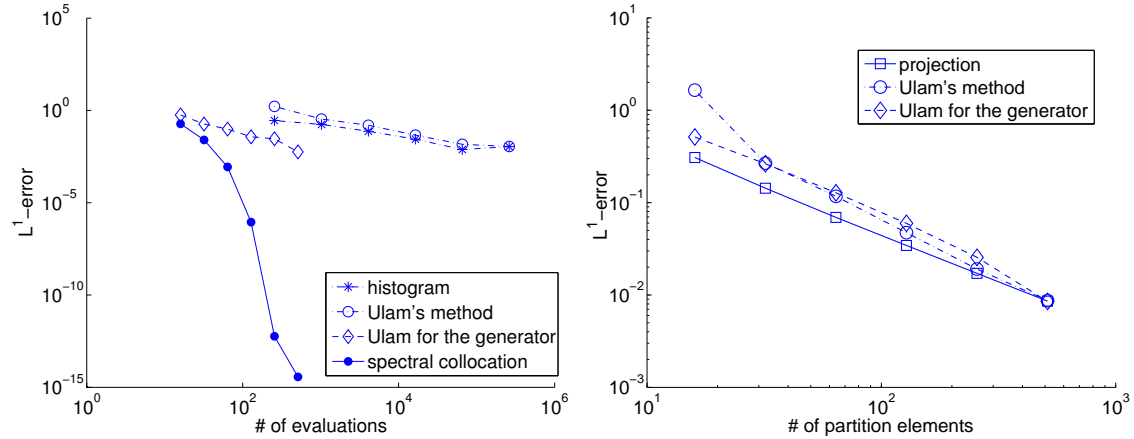

Figure 2: Left: L^1-error of the approximate invariant density as a function of the number of evaluations of the vector field (2^4 - 2^9 partition elements, resp. collocation points, 2^8 - 2^{18} number of iterations). Right: L^1-error of the approximate invariant densities obtained by projection of the true invariant density on the underlying partition, by Ulam's method, and by Ulam's method for the generator.

The number of grid sets (32) chosen in Figure 1 is tiny, and chosen merely for illustration purposes. Likewise, the maximum number of evaluations used in both generator schemes of 512 is also obviously tiny, and in practice one could very cheaply increase the number of grid sets and concomitantly the number of evaluations. The main message from this example is that in the right setting: low dimension, periodic domain, and infinitely smooth vector field, spectral collocation can significantly outperform standard Ulam and generator Ulam. Ulam's method for the generator is clearly outperforming standard Ulam in this example, and we will see that it continues to do so across a variety of dimensions, domains, and vector fields.

Table 1: Sources of computational cost. The dimension of the approximation space is denoted by n, d denotes the dimension of state space. In contrast to the other two methods, spectral collocation produces a full matrix. One can reduce the cost for solving the eigenproblem in this case to $\mathcal{O}(n\log(n)) \cdot \#$(GMRES iterations) by using an iterative solver (e.g. GMRES), and FFT to compute matrix-vector products.

method	approximation error	flops to set up matrix / time integration	dimension of spatial integrals	flops to solve EVP / vector iteration type
Ulam's method	$\mathcal{O}(n^{-1/d})$	$\mathcal{O}(n^2)$ / yes	d	$\mathcal{O}(n)$ / forw.
Ulam's method for the generator	$\mathcal{O}(n^{-1/d})$	$\mathcal{O}\big(n^{(2d-1)/d}\big)$ / no	$d-1$	$\mathcal{O}(n)$ / backw.
spectral collocation for the generator	$\mathcal{O}(n^{-k/d}),\ k \in \mathbb{N}$	$\mathcal{O}(n)$ / no	0	$\mathcal{O}(n^3)$ / backw.

5.2 An area-preserving cylinder flow

We consider an area-preserving flow on the cylinder, defined by interpolating a numerically given vector field as shown in Figure 3, which is a snapshot from a quasi-geostrophic flow, cf. [28]. The domain is periodic with respect to the x coordinate and the field is zero at the boundaries $y = 0$ and $y = 8 \cdot 10^5$. Again, we apply the three methods discussed in Section 4 in order to compute approximate eigenfunctions of the transfer operator and the generator.
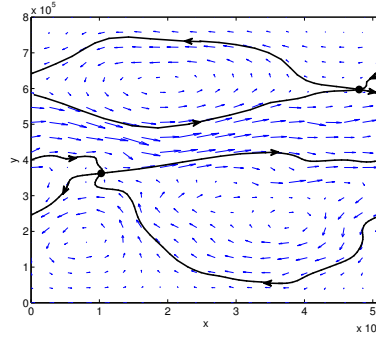

Figure 3: Vector field of the area-preserving cylinder flow. Shown are also two saddle equilibria (black dots) together with their stable and unstable manifolds.

5.2.1 Perturbing the model

Since the flow is area-preserving, we have a continuum of closed orbits, and each of them is an invariant set. On each, an invariant measure can be supported, i.e. the eigenspace at 1 (resp. 0) is infinite-dimensional. Both Ulam type methods can be interpreted as a small random perturbation of the original system (albeit different ones), yielding a unique stationary eigenvector of the resulting matrix. However, simultaneously with decreasing the cell size, the size of this "built-in" perturbation shrinks and it is therefore not immediate to which invariant measure the discretization schemes converge to. In order to make the problem well posed and the results comparable among the three methods we therefore need to artificially add a random perturbation (in form of a diffusion term) to the model.

We therefore choose the noise level ε such that the resulting diffusion coefficient $\varepsilon^2/2$ has the same order of magnitude as the numerical diffusion present within Ulam's method for the generator, but is still larger than the latter. Since the estimate from Remark 4.6 yields a numerical diffusion coefficient of ≈ 120 (see also [18]), we choose $\varepsilon = \sqrt{2 \cdot 500}$ here.

In **Ulam's method**, for the simulation of the SDE (3) a fourth order Runge-Kutta method is used, where in every time step a properly scaled (by a factor $\sqrt{h} \cdot \varepsilon$, where h is the time

step) normally distributed random number is added. We use 1000 sample points per box and the integration time $T = 5 \cdot 10^6$, which is realized by 20 steps of the Runge–Kutta method. Note that the integrator does not know that the flow lines should not cross the lower and upper boundaries of the state space. Points that leave phase space are projected back along the y axis into the next boundary box. An adaptive step–size control could resolve this problem, however at the cost of even more right hand side evaluations. The domain is partitioned into 128×128 boxes. For **Ulam's method for the generator**, again, we employ a partition of 128×128 boxes and approximate the edge integrals by the trapezoidal rule using three nodes (which turns out to be sufficiently accurate in this case). For **spectral collocation**, we employ 51 Fourier modes in the x coordinate (periodic boundary conditions) and the first 51 Chebyshev polynomials in the y coordinate, together with Neumann boundary conditions. Due to spectral convergence, this smaller number of modes already yields an accuracy in the invariant density comparable to the one of the other two approaches.

5.2.2 Computing almost invariant sets

In Table 2 we compare approximations for the three leading eigenvalues from our three schemes. If we assume that spectral collocation gives a very accurate result, Ulam's method should be $\mathcal{O}(n^{-1})$ from it. With $n = 128$ this matches well with the numbers in the table. Ulam's method for the generator generates additional numerical diffusion; this scales the corresponding eigenvalues. In

Table 2: Approximate eigenvalues.

method	λ_2	λ_3	λ_4
Ulam's method $(\log(\lambda_i)/T)$	$-1.64 \cdot 10^{-8}$	$-0.91 \cdot 10^{-7}$	$-1.06 \cdot 10^{-7}$
Ulam's method for the generator	$-1.98 \cdot 10^{-8}$	$-1.03 \cdot 10^{-7}$	$-1.19 \cdot 10^{-7}$
spectral collocation for the generator	$-1.65 \cdot 10^{-8}$	$-0.92 \cdot 10^{-7}$	$-1.05 \cdot 10^{-7}$

Figure 4 we compare the approximate eigenvectors of the second, third and fourth relevant eigenvalues of the transfer operator (resp. generator). Clearly, they all give the same qualitative picture. Yet, the number of evaluations of the vector field (and thus the associated cpu times) differ significantly, as shown in Table 3. The regions enclosed by the stable/unstable manifolds of the two

Table 3: Number of evaluations of the vector field and cpu times (on a 2.5 GHz Core Duo) for computing the approximate operator resp. generator and cpu times for computing the leading 4 eigenvalues/-vectors.

method	# of rhs evals	time matrix	time `eigs`
Ulam's method	$\approx 3 \cdot 10^8$	3020 sec.	1.0 sec.
Ulam's method for the generator	$\approx 3 \cdot 10^5$	143 sec.	0.8 sec.
spectral collocation for the generator	$\approx 3 \cdot 10^3$	0.4 sec.	3 sec.

saddle equilibria shown in Figure 3 are obviously invariant for the noiseless flow. The introduction of noise means that it is possible (but with a rather low probability) for stochastic trajectories to jump between these regions, making these regions almost-invariant sets (cf. Section 3.2). Figure 4 shows that these almost-invariant sets are highlighted in the 2nd, 3rd, and 4th eigenvectors as level sets at high gradient regions in the eigenvectors (in this example, near the yellow bands). The 2nd and 3rd eigenvectors highlight regions near different stable/unstable manifolds at the yellow bands and the 4th eigenvector shown in Figure 5 shows a combination. We refer the reader to [9, 16] for details on the connection between invariant manifolds and almost invariant sets. As one sees in Figure 5, the eigenvectors from Ulam's method for the generator are smoother compared to the ones from the standard Ulam method. Ulam's method for the generator appears to be more accurate, at least for this system with a small amount of noise. A possible explanation is that with Ulam's method for the generator we formally average the diffusion, while with the standard

(a) Ulam's method

(b) Ulam's method for the generator
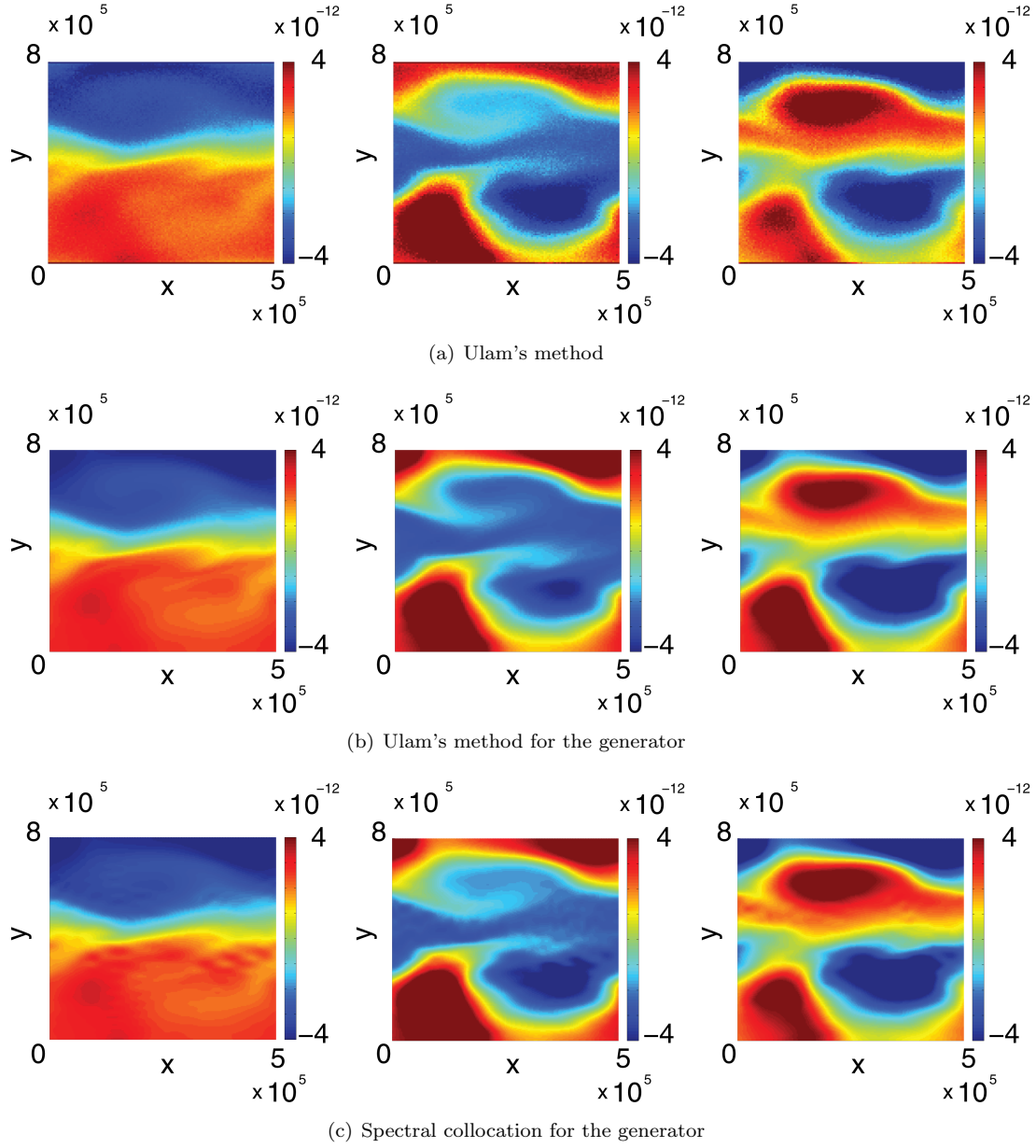

(c) Spectral collocation for the generator

Figure 4: Quasi-geostrophic flow: Eigenvectors at the second, third and fourth eigenvalue (from left to right).
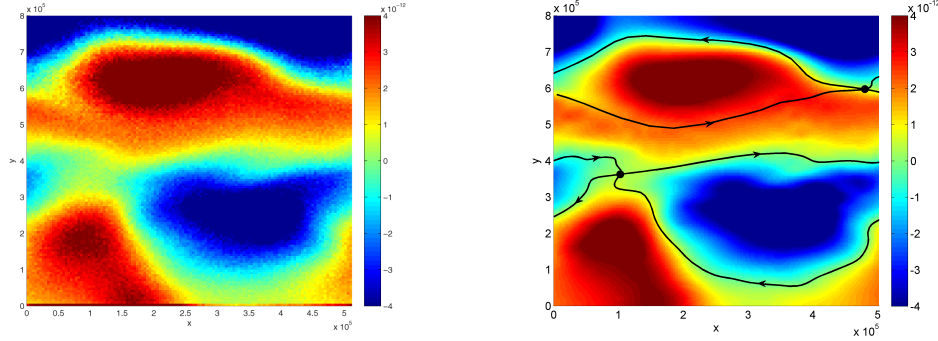
Figure 5: Comparison of the approximate eigenvector at λ_4 by Ulam's method (left) with the one computed by Ulam's method for the generator (right, together with the two saddle equilibria and their invariant manifolds).

Ulam method we simulate a finite number of random trajectories, which incompletely sample the true noise distribution.

5.3 A volume-preserving 3D example: the ABC-flow

In this section we demonstrate the efficacy of our generator approach for a three-dimensional flow. We consider the so-called ABC-flow [2], given by

$$
\begin{aligned}
\dot{x} &= a\sin(2\pi z) + c\cos(2\pi y) \\
\dot{y} &= b\sin(2\pi x) + a\cos(2\pi z) \\
\dot{z} &= c\sin(2\pi y) + b\cos(2\pi x),
\end{aligned}
$$

on the 3 dimensional torus. The flow is volume-preserving, so in order to make the computation of the spectrum well posed we need to add some artifical diffusion again. The vector field is very smooth, and in numerical experiments a diffusion coefficient $\varepsilon = 0.04$ turned out to work well. For $a = \sqrt{3}$, $b = \sqrt{2}$ and $c = 1$, there is numerical evidence that the flow exhibits complicated dynamics and invariant sets of complicated geometry [11, 16].

We approximate the leading few eigenfunctions of $\mathcal{A}_\varepsilon$ with Ulam's method for the generator on a $64 \times 64 \times 64$ grid. We used Gauss quadrature with 16 nodes on each face of a box in order to approximate the surface integrals, requiring $\sim 16 \cdot 64^3 \approx 4 \cdot 10^6$ evaluations of the vector field. The L^1-error of the approximate invariant density is $\approx 10^{-12}$. The second and fifth eigenfunctions are shown in Figure 6. The regions highlighted in red and blue are invariant cylinders of the deterministic flow; see [16] for more details on how to extract invariant and almost-invariant sets from the eigenfunctions.

We also apply spectral collocation on a $11 \times 11 \times 11$ grid, requiring $11^3 = 1331$ evaluations of the vector field. The L^1-error of the approximate invariant density is $\approx 10^{-14}$. We observe fast convergence (after comparing with computations on coarser resolutions) of not just the first nontrivial eigenfunctions, but also of higher order ones.
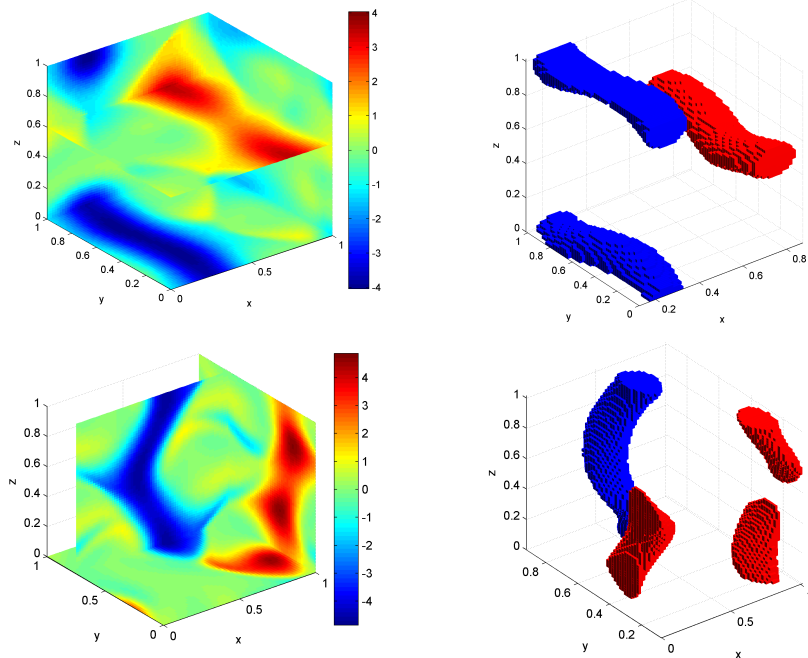
Figure 6: Second and fifth eigenvectors of the infinitesimal generator – Ulam type discretization. Left: slices, right: regions, where the absolute value of the left eigenfunctions exceed a given threshold – the "core" of the almost invariant sets, representing invariant cylinders in this example.

5.4 A dissipative 3D example with complicated geometry: the Lorenz system

As a last example we consider a system where the effective dynamics is supported on a set of complicated geometry and fractional dimension, the well-known Lorenz system

$$\dot{x} = \sigma(y - x),$$
$$\dot{y} = x(\varrho - z) - y,$$
$$\dot{z} = xy - \beta z,$$

where we make the standard parameter choices $\sigma = 10$, $\varrho = 28$ and $\beta = 8/3$. Since the effective (complicated) dynamics happens on a lower dimensional attractor we do not expect spectral collocation to work well and we therefore apply Ulam's method for the generator here.

A decade ago, numerical techniques were introduced for computing box coverings for attractors of complicated structure, cf. [7, 6]. These techniques make use of the fact that the set M to be computed is an attractor, hence each trajectory starting in its vicinity will be pulled to M in a fairly short time. In our approach time is not considered, we only use the vector field. Since the boundary of the box covering does not have to coincide with the boundary of the attractor, a tight box covering might not show the desired results, because of relatively big outflow rates in boundary boxes. The simplest idea is to use a big enough rectangle – in our case $[-30, 30] \times [-30, 30] \times [-10, 70]$. However, this set is not invariant, hence the question arises, which boundary conditions one should impose. Translating the condition "no probability flow on the boundary" in this Ulam type approach leads to the condition that the flow rates at the boundary are ignored. The attractor is then extracted by simple thresholding: regions where the approximate invariant density is nonzero are expected to lie in a small neighbourhood of the attractor. The finer the resolution, the smaller the diffusion introduced by the discretization and thus the tighter this neighbourhood of the attractor is. Having constructed the attractor, we may restrict the other eigenfunctions to this set in order to obtain almost invariant sets within the attractor.
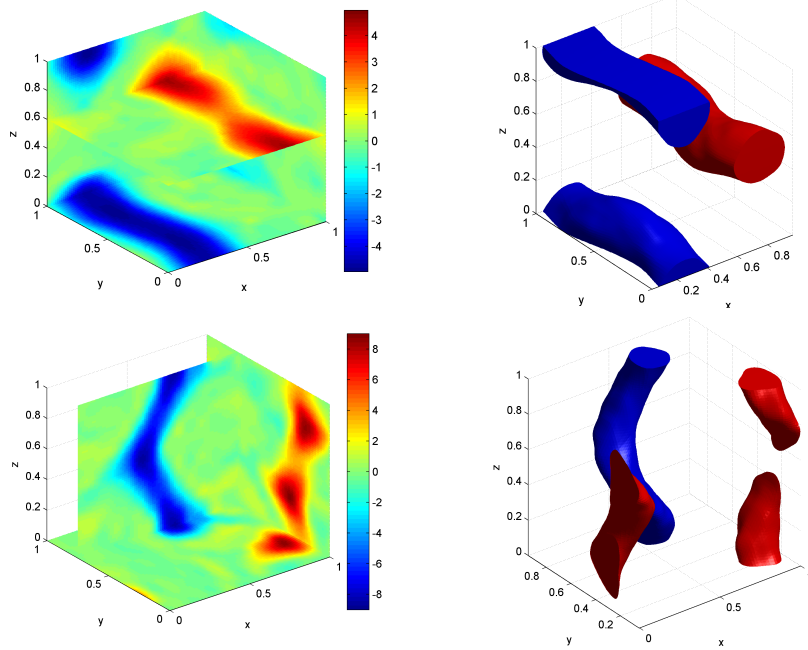
Figure 7: Second and fifth eigenfunction of the generator computed by spectral collocation. Left: slices, right: regions, where the absolute value of the left eigenfunctions exceed a given threshold – the "core" of the almost invariant sets.

For our computation, we employed a grid of $128 \times 128 \times 128$ boxes and used Gauss quadrature in 5×5 nodes on the faces in order to set up the approximate generator. The attractor is extracted by thresholding the approximate invariant density u: We cut off at a threshold value $5 \cdot 10^{-6}$ such that 96% of the invariant measure is supported on $\{u_1 > c\}$. Figure 8 shows the corresponding covering as well as the sign structure of the second and third eigenvector.
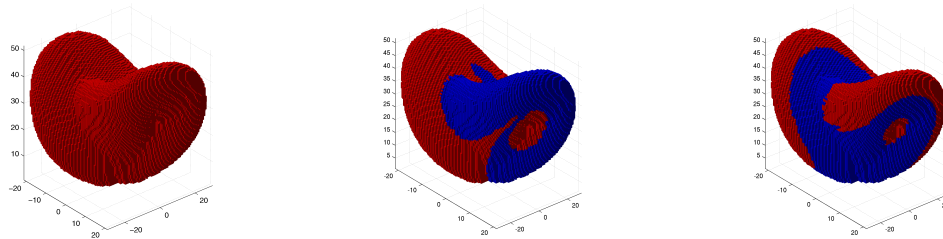

Figure 8: Ulam's method for the generator in an example with a low-dimensional attractor: The Lorenz attractor (left) and almost invariant sets: sign structure of the second eigenvector (middle) and third eigenvector (right).

6   Conclusion

We introduced the infinitesimal generator as a tool for studying long term behavior of flows without requiring any trajectory integration. The (typically unit dimensional) null space of the generator is spanned by the invariant density of the flow; the invariant density describes the asymptotic distribution of trajectories. We showed that eigenfunctions of the generator corresponding to

eigenvalues close to 0 had strong connections with almost-invariant or metastable behavior, and could be used to identify regions in the flow from which the escape rate is very low.

We proposed two new numerical approaches for the estimation of the generator; one based on Ulam's method and the other based on spectral collocation. We tested our numerical approaches on "full phase space" flows in one, two, and three dimensions and found that both approaches significantly outperformed the standard Ulam approach based on the Perron-Frobenius operator. Moreover, all operator/generator based methods were superior to standard trajectory integration for estimating the invariant density; the identification of almost-invariant sets is essentially impossible without an operator/generator approach. The spectral collocation approach worked particularly well in periodic domains when the vector field is infinitely smooth.

Acknowledgements

References

[1] H. AMANN, *Dual semigroups and second order linear elliptic boundary value problems*, Israel Journal of Mathematics, 45. (1983), pp. 225–254.

[2] V. ARNOL'D, *Sur la topologie des écoulements stationnaires des fluides parfaits*, C. R. Acad. Sci. Paris, 261 (1965), pp. 17–20.

[3] A. BERMAN AND R.J. PLEMMONS, *Nonnegative matrices in the mathematical sciences*, SIAM, 1994.

[4] J.P. BOYD, *Chebyshev and Fourier Spectral Methods*, Dover Publications, 2. ed., 2001.

[5] C. CANUTO, M.V. HUSSAINI, A. QUARTERONI, AND T.A. ZANG, *Spectral methods in fluid dynamics*, Springer Series in Computational Physics, Springer-Verlag, 3. ed., 1991.

[6] M. DELLNITZ, G. FROYLAND, AND O. JUNGE, *The algorithms behind GAIO–set oriented numerical methods for dynamical systems*, in Ergodic theory, analysis, and efficient simulation of dynamical systems, Bernold Fiedler, ed., Springer, Berlin, 2001, pp. 145–174, 805–807.

[7] M. DELLNITZ AND A,?HOHMANN, *A subdivision algorithm for the computation of unstable manifolds and global attractors*, Numer. Math., 75 (1997), pp. 293–317.

[8] M. DELLNITZ AND O. JUNGE, *On the approximation of complicated dynamical behavior*, SIAM J. Numer. Anal., 36 (1999), pp. 491–515.

[9] M. DELLNITZ, O. JUNGE, W.S. KOON, F. LEKIEN, M.W. LO, J.E. MARSDEN, K. PAD-BERG, R. PREIS, S.D. ROSS, AND B. THIERE, *Transport in dynamical astronomy and multi-body problems*, Int. J. Bifurcation Chaos Appl. Sci. Eng., 15 (2005), pp. 699–727.

[10] J. DING, Q. DU, AND T. Y. LI, *High order approximation of the Frobenius-Perron operator*, Appl. Math. Comp., 53 (1993), pp. 151–171.

[11] T. DOMBRE, U. FRISCH, M. HENON, J. M. GREENE, AND A. M. SOWARD, *Chaotic stream-lines in the ABC flows*, Journal of Fluid Mechanics, 167 (1986), pp. 353–391.

[12] G. FROYLAND, *Finite approximation of Sinai-Bowen-Ruelle measures for Anosov systems in two dimensions*, Random and Computational Dynamics, 3 (1995), pp. 251–264.

[13] G. FROYLAND, *Extracting dynamical behaviour via Markov models*, in Nonlinear dynamics and statistics, Alistair I. Mees, ed., Birkhauser, 2001, pp. 281–321.

[14] G. Froyland, *Statistically optimal almost-invariant sets*, Physica D, 200 (2005), pp. 205–219.

[15] G. Froyland and M. Dellnitz, *Detecting and locating near-optimal almost-invariant sets and cycles*, SIAM J. Sc. Comp., 24 (2003), pp. 1839–1863.

[16] G. Froyland and K. Padberg, *Almost-invariant sets and invariant manifolds – connecting probabilistic and geometric descriptions of coherent structures in flows.*, Physica D, 238 (2009), pp. 1507–1523.

[17] G. Froyland and O. Stancevic, *Escape rates and Perron-Frobenius operators: Open and closed dynamical systems*, Disc. Cont. Dyn. Sys. –Series B, 14 (2010), pp. 457–472.

[18] P. Koltai, *Efficient approximation methods for the global long-term behavior of dynamical systems – Theory, algorithms and examples*, PhD thesis, Technische Universität München, 2010.

[19] D. Kröner, *Numerical Schemes for Conservation Laws*, Wiley & Teubner, 1997.

[20] A. Lasota and M.C. Mackey, *Chaos, fractals, and noise*, vol. 97 of Applied Mathematical Sciences, Springer-Verlag, New York, second ed., 1994.

[21] R.J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press, 2002.

[22] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Birkhäuser, 1995.

[23] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, vol. 44 of Applied Mathematical Sciences, Springer-Verlag, New York, 1983.

[24] G. Pianigiani and J.A. Yorke, *Expanding maps on sets which are almost invariant: decay and chaos*, Transactions of the American Mathematical Society, (1979), pp. 351–366.

[25] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, no. 23 in Springer Series in Computational Mathematics, Springer, 1994.

[26] O. Stancevic, *Transfer operator methods in continuous time dynamical systems*. Honours thesis, University of New South Wales, 2007.

[27] L.N. Trefethen, *Spectral methods in MATLAB*, vol. 10 of Software, Environments, and Tools, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[28] A.M. Treguier and R.L. Panetta, *Multiple zonal jets in a quasigeostrophic model of the Antarctic Circumpolar Current*, J. Physical Oceanography, 24 (1994), pp. 2263–2277.

[29] L.-S. Young, *What are SRB measures, and which dynamical systems have them?*, J. Statist. Phys., 108 (2002), pp. 733–754.

[30] E.C. Zeeman, *Stability of dynamical systems*, Nonlinearity, 1 (1988), pp. 115–155.